

SPEECH RECOGNITION DEVICE

Publication number: JP10282987 (A)

Publication date: 1998-10-23

Inventor(s): WAKIZAKA SHINJI; ISHIWATARI KAZUYOSHI; KONDO KAZUO

Applicant(s): HITACHI LTD

Classification:

- international: **G09B29/10; G01C21/00; G01S5/14; G08G1/137; G10L15/00; G10L15/28; G09B29/10; G01C21/00; G01S5/14; G08G1/123; G10L15/00; (IPC1-7): G10L3/00; G01C21/00; G01S5/14; G08G1/137; G09B29/10; G10L3/00**

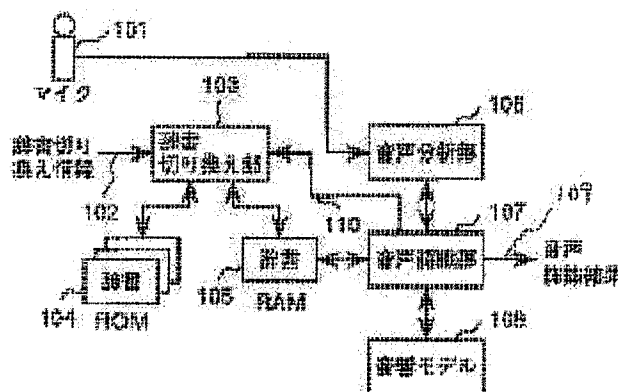
- European:

Application number: JP19970086344 19970404

Priority number(s): JP19970086344 19970404

Abstract of JP 10282987 (A)

PROBLEM TO BE SOLVED: To prevent performances of a recognition rate and a recognition response time from deteriorating and also to improve a user interface of a car navigation system even if vocabulary recognizable in speech is increased in the number as a whole system, in a speech recognition system used for a car navigation system, etc. **SOLUTION:** This system collects words and sentences for speech recognition and defines them as dictionaries 104, 105, and picks up those words and sentences as a result of speech recognition. In this case, this system is provided with a plurality of dictionaries 104, 105, and one of the plural dictionaries 104, 105 is selected by a dictionary change-over part 103, and speech recognition is executed by referring to it. A distance up to a destination, time, route, etc., are displayed on a map by using a result of the speech recognition in a car navigation system.



Data supplied from the **esp@cenet** database — Worldwide

(19)日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開平10-282987

(43)公開日 平成10年(1998)10月23日

(51)Int.Cl.⁶
G 1 0 L 3/00

G 0 1 C 21/00
G 0 1 S 5/14
G 0 8 G 1/137

識別記号
5 5 1
5 7 1

F I
G 1 0 L 3/00 5 5 1 J
5 7 1 C

G 0 1 C 21/00 H
G 0 1 S 5/14
G 0 8 G 1/137

審査請求 未請求 請求項の数19 OL (全 13 頁) 最終頁に続く

(21)出願番号 特願平9-86344

(22)出願日 平成9年(1997)4月4日

特許法第65条の2第2項第4号の規定により図面第10図の一部は不掲載とする。

(71)出願人 000005108
株式会社日立製作所
東京都千代田区神田駿河台四丁目6番地
(72)発明者 麻坂 新路
東京都小平市上水本町五丁目20番1号 株式会社日立製作所半導体事業部内
(72)発明者 石渡 一嘉
東京都小平市上水本町五丁目20番1号 株式会社日立製作所半導体事業部内
(72)発明者 近藤 和夫
東京都小平市上水本町五丁目20番1号 株式会社日立製作所半導体事業部内
(74)代理人 弁理士 高橋 明夫 (外1名)

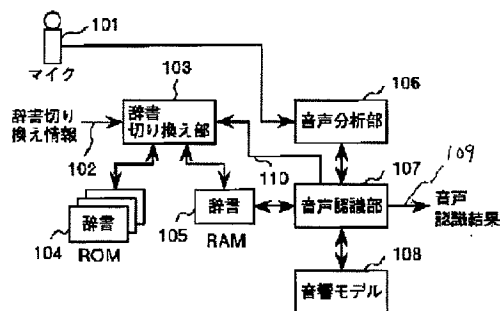
(54)【発明の名称】 音声認識システムおよび方法

(57)【要約】

【課題】カーナビゲーションシステムなどに用いられる音声認識システムにおいて、システム全体として音声認識できる語彙数の増加しても、認識率や認識応答時間の性能を低下させないようにする。また、カーナビゲーションシステムのユーザインターフェイスを向上させる。

【解決手段】音声認識の対象となる単語や文章を集めて辞書として定義し、音声認識の結果として、それらの単語や文章をピックアップする音声認識システムにおいて、辞書を複数持たせ、辞書切り換え部により、複数の辞書より一つの辞書を選択して、それを音声認識の対象として、音声認識をおこなう。音声認識の結果を用いて、カーナビゲーションシステム際に地図上に、目的地までの距離、時間、ルートなどを表示する。

図 1



【特許請求の範囲】

【請求項1】 音声認識の対象となる単語や文章を集めて辞書として定義し、音声認識の結果として、それらの単語や文章をピックアップする音声認識システムにおいて、

このシステムは、

前記辞書を複数有し、

前記複数の辞書を格納しておく第一の記憶部と、

前記複数の辞書から一つの辞書を選択し格納しておく第二の記憶部と、

前記複数の辞書から一つだけ辞書を選択する辞書切り換え情報を受けて、辞書を切り換える辞書切り換え部と、取り込んだ音声に対して、音声分析処理をおこなう音声分析部と、

音声のパターンを音素単位で捉える音響モデルと、

音声分析結果に対して、前記音響モデルと前記辞書とを参照し、音声認識処理をおこなう音声認識部とを備え、前記辞書切り換え部により、前記複数の辞書より一つの辞書を選択して、それを音声認識の対象として、音声認識をおこなうことを特徴とする音声認識システム。

【請求項2】 前記辞書の切り換え情報が、音声認識の結果、音声認識の対象とした辞書の中に該当する単語や文章が見出せなかったことであることを特徴とする請求項1記載の音声認識システム。

【請求項3】 前記辞書に集められた単語や文章に、その単語や文章を表すコードの外に、付加情報を有することを特徴とする請求項1および請求項2記載のいずれかの音声認識システム。

【請求項4】 前記複数の辞書を格納しておく第一の記憶部は、ハードディスク、メモリカード、または、ROMであり、

前記複数の辞書から一つの辞書を選択し格納しておく第二の記憶部は、RAMであることを特徴とする請求項1ないし請求項3記載のいずれかの音声認識システム。

【請求項5】 この音声認識システムは、カーナビゲーションシステムにおける音声認識システムであって、前記辞書に集められた単語は、このカーナビゲーションシステムで用いられる地名、交差点名、通り名、建物名であることを特徴とする請求項1ないし請求項4記載のいずれかの音声認識システム。

【請求項6】 前記カーナビゲーションシステムは、カーナビゲーションの対象となるエリアを複数持ち、前記辞書は、各エリアに対応して設けられ、その辞書に集められた単語は、対応するエリアに存在する対象を表す地名、交差点名、通り名、建物名であることを特徴とする請求項5記載の音声認識システム。

【請求項7】 前記複数の辞書から一つの辞書を選択する辞書切り換え情報は、カーナビゲーションシステムで用いられている衛星測位システムGPS (Global Positioning system) からの位置情報であることを特徴とす

る請求項1ないし請求項6記載のいずれかの音声認識システム。

【請求項8】 この音声認識システムは、カーナビゲーションシステムにおける音声認識システムであって、前記辞書に集められた単語には、付加情報として、地球上の経度、緯度で示される位置情報を持ち、このカーナビゲーションシステムは、現在の走行するエリアを地図として表示しているときに、

現在地を示すGPSからの位置情報と、単語に付加された位置情報とから、音声認識した単語にあたる対象の前記地図上の表示位置座標値と、現在地からの走行距離とを計算して、

音声認識した単語にあたる対象の位置を地図上に表示するとともに、走行ルート、走行距離、走行所要時間を前記地図上に表示することを特徴とする請求項7記載の音声認識システム。

【請求項9】 第一の辞書に集められた単語が、第二の辞書に切り換えるためのインデックスであるインデックス辞書であり、

先ず、第一の辞書を参照して、音声認識をおこなって、その結果としてインデックスをピックアップし、そのインデックスに基づき、次の音声認識で参照する第二の辞書を選択することを特徴とする請求項1ないし請求項4記載のいずれかの音声認識システム。

【請求項10】 この音声認識システムは、携帯型情報機器における音声認識システムであって、

第一の辞書のインデックスが、

その携帯型情報機器の機能を表す単語、または、その携帯型情報機器に与えるコマンドを表す単語であることを特徴とする請求項9記載の音声認識システム。

【請求項11】 音声認識の対象となる単語や文章を集めて辞書として定義し、音声認識の結果として、それらの単語や文章をピックアップする音声認識システムを用いる音声認識方法において、

このシステムは、

前記辞書を複数有し、

前記複数の辞書を格納しておく第一の記憶部と、

前記複数の辞書から一つの辞書を選択し格納しておく第二の記憶部と、

前記複数の辞書から一つだけ辞書を選択する辞書切り換え情報を受けて、辞書を切り換える辞書切り換え部と、取り込んだ音声に対して、音声分析処理をおこなう音声分析部と、

音声のパターンを音素単位で捉える音響モデルと、

音声分析結果に対して、前記音響モデルと前記辞書とを参照し、音声認識処理をおこなう音声認識部とを備え、前記辞書切り換え部により、前記複数の辞書より一つの辞書を選択して、それを音声認識の対象として、音声認識をおこなうことを特徴とする音声認識方法。

【請求項12】 前記辞書の切り換え情報が、音声認識

の結果、音声認識の対象とした辞書の中に該当する単語や文章が見出せなかったことであることを特徴とする請求項1記載の音声認識方法。

【請求項13】 前記辞書に集められた単語や文章に、その単語や文章を表すコードの外に、付加情報を有することを特徴とする請求項11および請求項12記載のいずれかの音声認識方法。

【請求項14】 この音声認識システムは、カーナビゲーションシステムにおける音声認識システムであって、前記辞書に集められた単語は、このカーナビゲーションシステムで用いられる地名、交差点名、通り名、建物名であることを特徴とする請求項11ないし請求項13記載のいずれかの音声認識方法。

【請求項15】 前記カーナビゲーションシステムは、カーナビゲーションの対象となるエリアを複数持ち、前記辞書は、各エリアに対応して設けられ、その辞書に集められた単語は、対応するエリアに存在する対象を表す地名、交差点名、通り名、建物名であることを特徴とする請求項14記載の音声認識方法。

【請求項16】 前記複数の辞書から一つの辞書を選択する辞書切り換え情報は、カーナビゲーションシステムで用いられている衛星測位システムGPS (Global Positioning system) からの位置情報であることを特徴とする請求項11ないし請求項15記載のいずれかの音声認識方法。

【請求項17】 この音声認識方法は、カーナビゲーションシステムにおける音声認識方法であって、前記辞書に集められた単語には、付加情報として、地球上の経度、緯度で示される位置情報を持ち、このカーナビゲーションシステムは、現在の走行するエリアを地図として表示しているときに、現在地を示すGPSからの位置情報と、単語に付加された位置情報とから、音声認識した単語にあたる対象の前記地図上の表示位置座標値と、現在地からの走行距離とを計算して、音声認識した単語にあたる対象の位置を地図上に表示するとともに、走行ルート、走行距離、走行所要時間を前記地図上に表示することを特徴とする請求項16記載の音声認識方法。

【請求項18】 第一の辞書に集められた単語が、第二の辞書に切り換えるためのインデックスであるインデックス辞書であり、先ず、第一の辞書を参照して、音声認識をおこなって、その結果としてインデックスをピックアップし、そのインデックスに基づき、次の音声認識で参照する第二の辞書を選択することを特徴とする請求項11ないし請求項13記載のいずれかの音声認識方法。

【請求項19】 この音声認識方法は、携帯型情報機器における音声認識方法であって、第一の辞書のインデックスが、

その携帯型情報機器の機能を表す単語、または、その携帯型情報機器に与えるコマンドを表す単語であることを特徴とする請求項18記載の音声認識方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、音声認識システムおよび方法に係り、カーナビゲーションシステム、PDAに代表される小型情報機器、携帯型音声翻訳機に用いる音声認識システムであって、特に、カーナビゲーションシステムなどで、地名、交差点名、通り名等により、目的地等の音声検索、音声探索、音声認識誘導をおこなうような膨大な単語の音声認識に用いて好適な音声認識システムおよび方法に関する。

【0002】

【従来の技術】近年、音声認識技術を用いた小型情報システムが普及しつつある。カーナビゲーションシステムをはじめとして、PDAに代表される携帯型情報機器、携帯型翻訳機等である。

【0003】このような音声認識システムの例として、特開平5-35776号公報の「言語自動選択機能付翻訳装置」には、マイクから入力した操作者の音声进行認識して、翻訳し、翻訳した言語の音声を出力するようにした携帯用の翻訳装置に関する技術が開示されている。

【0004】以下、図11を用いてこのような従来技術に係る音声翻訳装置の概要について説明しよう。図11は、従来技術に係る音声翻訳装置の構成を示すブロック図である。

【0005】制御部801は、マイクロプロセッサ等からなり、装置の各部を制御する。音声区間切出し部802は、マイク809から入力された音声をデジタル信号に変換して切り出し、音声認識部803に送る。音声認識部803は、キーボード又はスイッチ等による操作信号811を受けた制御部801の指示により、マイク809、音声区間切出し部802を経て、切り出された音声を分析する。そして、その結果を、音声認識辞書部807に格納された標準音声パターンと比較することにより、音声認識をおこなう。音声合成部805は、音声認識部803により認識された音声に対応した翻訳語を、翻訳語データ用メモ리카ード806から読み込み、音声信号に変換してスピーカアンプ810、スピーカ808を経て出力する。

【0006】表示部804は、翻訳装置の使用者への指示や翻訳語の文字による表示等をおこなう。翻訳語データ用メモ리카ード806は、ROMカード等からなり、翻訳語を音声合成して出力する場合には、音声データを格納している。また、この翻訳語データ用メモ리카ード806から、翻訳語に対応したキャラクターコードを読み込み、表示部804に表示する。そして、この翻訳語データ用メモ리카ード806を他の言語のものと交換することにより、複数の言語に翻訳することが可能とな

る。音声認識辞書部807は、RAM等からなり、操作者の発生に応じた標準音声パターンを格納している。この標準音声パターンは、操作者があらかじめ格納しておく。

【0007】

【発明が解決しようとする課題】このような音声認識技術の分野は、半導体技術の向上を背景として、システムがより人間的なユーザインターフェイスを提供すべきであるという要望から、その発展が期待されている。上記従来の音声認識技術を用いた小型情報システムも、カーナビゲーションシステムをはじめとして、PDAに代表される携帯型情報機器、携帯型翻訳機として、今後ますます普及してくることが予想される。

【0008】しかしながら、音声認識は、処理すべき情報量が膨大なものになるため、従来の技術では、認識率や認識応答時間の性能を低下させないためには、認識する語数に制約を設ける必要がある。というのも、音声認識を用いたヒューマンインターフェイスの向上においては、認識率、および認識応答時間が問題となるからである。

【0009】従来の技術では、認識率や認識応答時間の性能を低下させないために、認識する語数を制約しなければならない。認識する語数を増やすと、音声の特徴が似通った単語が増加して認識率が低下する。また、認識対象となるすべての単語に対して、音声認識処理をおこなうので、そのために必要なワークメモリや辞書メモリ等の規模が大きくなり、処理ステップも増え、処理時間が増加する。

【0010】今後、音声認識技術の革新や、それを実現するソフトウェア、ハードウェアの性能向上により、認識する語数の制約が緩和されることも考えられるが、当面は、認識率や認識応答時間の性能を低下させないために、認識する語数を制約せざるを得ないという問題点がある。

【0011】その反面、カーナビゲーションシステムなどの音声認識技術を用いた小型情報システムでは、使い勝手を良くするために、音声認識する語数数の増加したいという要望がある。

【0012】また、従来のカーナビゲーションシステムは、目的地を入力して、ルートを表示する機能や、現在の地点、目的地までの距離を音声認識により問い合わせる機能はあるが、これらを有機的に結合して良好なユーザインターフェイスを提供した技術は、知られていない。

【0013】本発明は、上記問題点を解決するためになされたもので、その目的は、小型情報システムに用いられる音声認識システムにおいて、システム全体として音声認識できる語数数の増加しても、認識率や認識応答時間の性能を低下させないで音声認識ができる音声認識システムを提供することである。

【0014】また、本発明の今一つの目的は、音声認識を用いたカーナビゲーションシステムにおいて、特に、音声による目的地の地図上の位置の探索や、音声による誘導システムにおいて、良好な音声認識インターフェースを実現することである。

【0015】

【課題を解決するための手段】上記目的を達成するために、本発明の音声認識システムに係る発明の第一の構成は、音声認識の対象となる単語や文章を集めて辞書として定義し、音声認識の結果として、それらの単語や文章をピックアップする音声認識システムにおいて、このシステムは、前記辞書を複数有し、前記複数の辞書を格納しておく第一の記憶部と、前記複数の辞書から一つの辞書を選択し格納しておく第二の記憶部と、前記複数の辞書から一つだけ辞書を選択する辞書切り換え情報を受けて、辞書を切り換える辞書切り換え部と、取り込んだ音声に対して、音声分析処理をおこなう音声分析部と、音声のパターンを音素単位で捉える音響モデルと、音声分析結果に対して、前記音響モデルと前記辞書とを参照し、音声認識処理をおこなう音声認識部とを備え、前記辞書切り換え部により、前記複数の辞書より一つの辞書を選択して、それを音声認識の対象として、音声認識をおこなうようにしたものである。

【0016】より詳しくは、上記音声認識システムにおいて、前記辞書の切り換え情報が、音声認識の結果、音声認識の対象とした辞書の中に該当する単語や文章が見出せなかったことであるようにしたものである。

【0017】また詳しくは、上記音声認識システムにおいて、前記辞書に集められた単語や文章に、その単語や文章を表すコードの外に、付加情報を有するようにしたものである。

【0018】さらに詳しくは、上記音声認識システムにおいて、前記複数の辞書を格納しておく第一の記憶部は、ハードディスク、メモリカード、または、ROMであり、前記複数の辞書から一つの辞書を選択し格納しておく第二の記憶部は、RAMであるようにしたものである。

【0019】上記目的を達成するために、本発明の音声認識システムに係る発明の他の構成は、上記音声認識システムにおいて、この音声認識システムは、カーナビゲーションシステムにおける音声認識システムであって、前記辞書に集められた単語は、このカーナビゲーションシステムで用いられる地名、交差点名、通り名、建物名であるようにしたものである。

【0020】より詳しくは、前記カーナビゲーションシステムは、カーナビゲーションの対象となるエリアを複数持ち、前記辞書は、各エリアに対応して設けられ、その辞書に集められた単語は、対応するエリアに存在する対象を表す地名、交差点名、通り名、建物名であるようにしたものである。

【0021】また詳しくは、上記音声認識システムにおいて、前記複数の辞書から一つの辞書を選択する辞書切り換え情報は、カーナビゲーションシステムで用いられている衛星測位システムGPS (Global Positioning system) からの位置情報であるようにしたものである。

【0022】さらに詳しくは、上記音声認識システムにおいて、この音声認識システムは、カーナビゲーションシステムにおける音声認識システムであって、前記辞書に集められた単語には、付加情報として、地球上の経度、緯度で示される位置情報を持ち、このカーナビゲーションシステムは、現在の走行するエリアを地図として表示しているときに、現在地を示すGPSからの位置情報と、単語に付加された位置情報とから、音声認識した単語にあたる対象の前記地図上の表示位置座標値と、現在地からの走行距離とを計算して、音声認識した単語にあたる対象の位置を地図上に表示するとともに、走行ルート、走行距離、走行所要時間を前記地図上に表示するようにしたものである。

【0023】上記目的を達成するために、本発明の音声認識システムに係る発明のまた他の構成は、第一の辞書に集められた単語が、第二の辞書に切り換えるためのインデックスであるインデックス辞書であり、先ず、第一の辞書を参照して、音声認識をおこなって、その結果としてインデックスをピックアップし、そのインデックスに基づき、次の音声認識で参照する第二の辞書を選択するようにしたものである。

【0024】より詳しくは、上記音声認識システムにおいて、この音声認識システムは、携帯型情報機器における音声認識システムであって、第一の辞書のインデックスが、その携帯型情報機器の機能を表す単語、または、その携帯型情報機器に与えるコマンドを表す単語であるようにしたものである。

【0025】上記目的を達成するために、本発明の音声認識方法に係る発明の構成は、音声認識の対象となる単語や文章を集めて辞書として定義し、音声認識の結果として、それらの単語や文章をピックアップする音声認識システムを用いる音声認識方法において、このシステムは、前記辞書を複数有し、前記複数の辞書を格納しておく第一の記憶部と、前記複数の辞書から一つの辞書を選択し格納しておく第二の記憶部と、前記複数の辞書から一つだけ辞書を選択する辞書切り換え情報を受けて、辞書を切り換える辞書切り換え部と、取り込んだ音声に対して、音声分析処理をおこなう音声分析部と、音声のパターンを音素単位で捉える音響モデルと、音声分析結果に対して、前記音響モデルと前記辞書とを参照し、音声認識処理をおこなう音声認識部とを備え、前記辞書切り換え部により、前記複数の辞書より一つの辞書を選択して、それを音声認識の対象として、音声認識をおこなうようにしたものである。

【0026】より詳しくは、上記音声認識方法におい

て、前記辞書の切り換え情報が、音声認識の結果、音声認識の対象とした辞書の中に該当する単語や文章が見出せなかったことであるようにしたものである。

【0027】また詳しくは、上記音声認識方法において、前記辞書に集められた単語や文章に、その単語や文章を表すコードの外に、付加情報を有するようにしたものである。

【0028】上記目的を達成するために、本発明の音声認識方法に係る発明の他の構成は、この音声認識システムは、カーナビゲーションシステムにおける音声認識システムであって、前記辞書に集められた単語は、このカーナビゲーションシステムで用いられる地名、交差点名、通り名、建物名であるようにしたものである。

【0029】より詳しくは、上記音声認識方法において、前記カーナビゲーションシステムは、カーナビゲーションの対象となるエリアを複数持ち、前記辞書は、各エリアに対応して設けられ、その辞書に集められた単語は、対応するエリアに存在する対象を表す地名、交差点名、通り名、建物名であるようにしたものである。

【0030】また詳しくは、上記音声認識方法において、前記複数の辞書から一つの辞書を選択する辞書切り換え情報は、カーナビゲーションシステムで用いられている衛星測位システムGPS (Global Positioning system) からの位置情報であるようにしたものである。

【0031】さらに詳しくは、上記音声認識方法において、この音声認識方法は、カーナビゲーションシステムにおける音声認識方法であって、前記辞書に集められた単語には、付加情報として、地球上の経度、緯度で示される位置情報を持ち、このカーナビゲーションシステムは、現在の走行するエリアを地図として表示しているときに、現在地を示すGPSからの位置情報と、単語に付加された位置情報とから、音声認識した単語にあたる対象の前記地図上の表示位置座標値と、現在地からの走行距離とを計算して、音声認識した単語にあたる対象の位置を地図上に表示するとともに、走行ルート、走行距離、走行所要時間を前記地図上に表示するようにしたものである。

【0032】上記目的を達成するために、本発明の音声認識方法に係る発明のまた他の構成は、第一の辞書に集められた単語が、第二の辞書に切り換えるためのインデックスであるインデックス辞書であり、先ず、第一の辞書を参照して、音声認識をおこなって、その結果としてインデックスをピックアップし、そのインデックスに基づき、次の音声認識で参照する第二の辞書を選択するようにしたものである。

【0033】より詳しくは、上記音声認識方法において、この音声認識方法は、携帯型情報機器における音声認識方法であって、第一の辞書のインデックスが、その携帯型情報機器の機能を表す単語、または、その携帯型情報機器に与えるコマンドを表す単語であるようにした

10

20

30

40

50

ものである。

【0034】

【発明の実施の形態】以下、本発明に係る各実施形態を、図1ないし図10を用いて説明する。

【本発明の音声認識システムのシステム構成】先ず、図1および図2を用いて本発明の音声認識システムのシステム構成について説明する。図1は、本発明に係る音声認識システムの各機能とその処理の流れを示すブロック図である。図2は、本発明のハードウェア構成を示すブロック図である。

【0035】先ず、音声認識をおこなうために、図1に示されるマイク101から音声を取り込まれる。取り込まれた音声は、音声分析部106によってノイズ処理や音声分析などの前処理がなされ、音声認識部107により音声認識がなされる。ここで、音声認識とは、音声信号を解析して、それを音素に分析してそのパターンを解析し、該当する単語や文章を辞書から選択することである。そして、システムの出力として、音声認識結果109を生み出す。

【0036】音声認識部107は、音声分析部106で分析された入力音声の音声分析結果に対して、逐次、辞書105、および音響モデル108とを参照して、入力音声の照合をおこない、辞書105の中で、一番近い単語をピックアップする。

【0037】また、音響モデル108は、音声認識に用いられるモデルであり、具体的には、辞書に用いられている文字と音素との対応、また、音素の特徴を記憶したものである。音響モデルは、最近では、あらかじめ声を登録しなくても、誰が話し手もその声を認識できるいわゆる「不特定話者対応」が、一般的になってきている。このような音響モデルとしては、例えば、隠れマルコフモデル(HMM: Hidden Markov Model)を用いることができる。

【0038】音声認識部107では、音声認識のための辞書が用いられる。音声認識のために用いられる辞書とは、言葉、単語(名詞、動詞等)、文章を集めたものである。例えば、カーナビゲーションシステムにおいては、通り名、地名、建造物名、町名、番地、交差点名、個人住宅(個人名)等や、必要最小限の会話に必要な言葉の集合体である。より具体的には、運転している者が、発声する「ガソリンスタンド」、「コンビニエンスストア」、「ファミリーレストラン」等の言葉である。この辞書は、システムの能力に応じて、一つの辞書あたり、例えば、1000~5000語の単語で構成する。この辞書を複数用意して、音声認識の対象として、複数の辞書から一つの辞書を選択して音声認識をおこなう。

【0039】すなわち、音声認識は、音声分析結果より、音声モデルよりその特徴を参照して、当てはまる音素を見出し、その音素の並びより当てはまる語を辞書から検索し、ピックアップする処理である。

【0040】さて、本発明は、複数の辞書を持ち、辞書切り換え情報102を参照して、それを辞書切り換え部103で適宜切換えて、音声認識をおこなうのが特徴である。

【0041】辞書切り換え部103は、辞書切り換え情報102の内容にしたがって、音声認識の候補として、複数の辞書から一つの辞書を選択するか、または、切り換えるものである。例えば、複数の辞書がフラッシュメモリで構成されたメモリカードやROM(Read Only Memory)に格納されていて、音声認識するときに必要な辞書だけを、RAM(Random Access Memory)に転送して音声認識処理をおこなう。

【0042】このようにしたときには、この複数の辞書を置くための記憶装置(または、記憶領域)104は、メモリカードやROMで構成し、複数の辞書から一つの辞書を選択して格納するための記憶装置(または、記憶領域)105は、RAMで構成することになる。また、複数の辞書を格納しておくために、ハードディスクなどの補助記憶装置も用いることができる。

【0043】また、音声認識の結果は、信号110によって、音声認識部107から、辞書切り換え部103へフィードバックされる。これは、例えば、後にも説明するが、最適の単語が見つからないときには、別の辞書に切り換えることが考えられる。

【0044】なお、図1に示す各処理ブロックは、複数のLSIやメモリで構成されたシステムであっても、半導体素子上に構成された一つないし複数のシステムオンチップであってもよい。

【0045】次に、図2を用いて本発明に係る音声認識システムのハードウェア構成について説明する。

【0046】音声を取り込むためのマイク701は、カーナビゲーションシステム等では、周囲の雑音を取り込まないために指向性をもたせた指向性マイクである。

【0047】辞書を切り換えるためのデータ702(または、制御信号)は、カーナビゲーションシステムでは、GPSから送られてくる位置データである。

【0048】CPU703は、カーナビゲーションシステムや、PDA等のメインシステムの制御と、音声認識システムにおける音声認識処理をおこなう。このCPUには、RISCマイコンが用いられるのが、最近の潮流である。

【0049】A/D変換IC704は、マイク701により取り込まれたアナログ音声データをデジタル音声データに変換するチップである。

【0050】インターフェイス705は、辞書切り換えデータ702を受けて、CPU703に対して、辞書切り換え情報を読み込ませるためのインターフェースである。

【0051】ROM706は、辞書や音響モデル、プログラムを格納しておく記憶装置である。また、複数の辞

書を格納しておくために、メモリカードを用いても良い。

【0052】RAM707は、ROM706から転送された一部の辞書や、音響モデル、プログラムが格納され、また、音声認識処理に必要な必要最小限のワークメモリであり、ROM706に比べて、通常アクセス時間の短い半導体素子が用いられる。

【0053】バス708は、システムにおけるデータバス、アドレスバス、制御信号バスとして用いられる。

【0054】このようなハードウェア構成において、マイク701から取り込まれた音声は、辞書切り換えデータ702により、切り換えられた辞書を参照して、音声認識されることになる。辞書の切り換えは、CPU703がおこない、ROM706の全体の辞書の中から、必要に応じて、一部の辞書がRAM707へ転送される。そして、CPU703と、RAM707の間でデータ転送をしながら、一連の音声認識処理が進められることになる。

【0055】〔実施形態1〕以下、本発明に係る第一の実施形態を、図3ないし図7を用いて説明する。本実施形態では、本発明の音声認識システムをカーナビゲーションシステムに適用した場合について説明することにする。

(I) カーナビゲーションシステムの辞書切り換えについて

まず、図3を用いてカーナビゲーションシステムの辞書切り換えの具体的なイメージについて説明しよう。図3は、カーナビゲーションのエリアと対応する辞書の関係を説明するための模式図である。

【0056】本発明では、音声認識のための辞書を複数持ち、それを状況に応じて切り換えていくものである。

【0057】本実施形態のカーナビゲーションに用いられる音声認識システムは、走行するエリアに対応して音声認識の辞書を持つことにする。すなわち、このようにすれば、現在走行しているエリアに対する運転者の指示が有効におこなう事ができるからである。

【0058】このときに、辞書を切り換える条件としては、例えば、車がA地点からB地点まで走行したときに、A地点とB地点の距離がある一定の距離以上になったことが考えられる。本実施形態の音声認識システムは、上の状況においては、A地点で音声認識に使用していた辞書1から、B地点で音声認識に使用する辞書2に切り換えることになる。

【0059】以下、この例を図3を用いて詳細に説明しよう。図3(a)は、実際に、カーナビゲーションシステムを搭載した車が走行していく様子とエリアの関係を模式的に示した図である。この図で、丸で示した記号が車であり、それが道路301に沿って走行する。

【0060】エリア1の中の記号302は、カーナビゲーションシステムを搭載した車が現在走行しているポイ

ント(A地点)と走行方向を表示している。

【0061】A地点において、この音声認識システムが音声認識可能な単語は、矩形304が示すエリア1の中に存在する地名、通り名、交差点名、建造物名である。

【0062】ところでここで、表示されている縮尺度によって、エリアの中に存在する地名、通り名、交差点名、建造物名等の数は異なることは注意を要する。また、表示しているエリアが、市街地である場合と、田舎や山間部等の過疎地帯である場合とでも、エリアの中に存在する地名、通り名、交差点名、建造物名等の数は異なる。

【0063】そこで、カーナビゲーションシステムの地図の縮尺度 $1/k$ の k が大きい場合には、広範囲なエリアを表示していることになるので、単語数は増えることになる。例えば、音声認識において、認識率と認識応答時間の性能を低下させない単語数が、最大3000語とすると、3000語単位にエリアを分割することになる。反面、広範囲のエリアで音声認識をおこなうときには、大きな通り名や交差点名、有名な建造物名の単語で辞書を構成して、音声認識が実効するように単語を選択しなければならない。

【0064】逆に、縮尺度 $1/k$ の k が小さい場合には、狭い範囲のエリアを表示していることから、広いエリアを表示しているときと比べて単語数は減少する。しかしながら、縮尺度 $1/k$ の k が小さい場合にも、運転者は、より詳細な通り名や交差点名、建造物名を知りたいことから、できるだけ細かい通り名や交差点名、ローカルな建造物名まで含めて、単語数を増大させ、辞書の単語数は、音声認識の限界である3000語まで使用して辞書を構成することが望ましい。

【0065】ここで、音声認識のための辞書の語彙としては、エリア内に存在する建造物(ガソリンスタンド、コンビニエンスストア、レストラン等)について多様な検索ができるようにしておけば、ユーザの使い勝手が向上する。例えば、ガソリンスタンドについては、「ガソリンスタンド」という言葉自体、その供給メーカー名、そのガソリンスタンドの固有名詞である店の名前という具合である。

【0066】このようにしておけば、表示されているエリア1において、運転者が、カーナビゲーションシステムに対して、例えば「A社」と発声すると(ここで、A社はある特定のガソリン供給メーカを指すものとする)、エリア1内にA社系のガソリンスタンドが5km先に存在すれば、「5km先にあります。」と音声合成で答えてくれるようなユーザインターフェイスが提供することができる。

【0067】次に、A地点を走行していた車は、現在B地点を走行しているものとする。

【0068】記号303は、カーナビゲーションシステムを搭載した車が現在走行しているポイント(B地点)

10

20

30

40

50

と走行方向を表示している。B地点において、音声認識可能な単語は、矩形305が示すエリア2の中に存在する地名、通り名、交差点名、建造物名等である。

【0069】このように本実施形態の音声認識システムは、図3(b)に示されるように、エリアと辞書の関係を示すテーブル306を持っている。

【0070】各辞書は、対応するエリアの中に存在する地名、通り名、交差点名、建造物名、等の単語で構成されている。

【0071】(2) 辞書切り換えの処理

次に、図4を用いて音声認識の辞書の切り換え処理について説明しよう。図4は、本発明の第一の実施形態に係る音声認識システムにおいて、音声認識の辞書切り換え処理を示すフローチャートである。

【0072】本発明では、音声認識のための辞書を複数持ち、それを状況に応じて切り換えていく。

【0073】このときに、音声認識処理の前に、辞書切り換え情報が更新されたか否かを判定する(S501)。

【0074】辞書切り換え情報は、カーナビゲーションシステムであれば、衛星測位システム(GPS: Global Positioning System)からの位置を示す信号である。

【0075】図1に示された辞書切り換え部103は、GPSからの位置を示す信号を受けて、その位置が認識対象の単語辞書を切り換える必要がある事を示している場合(YES)には、認識対象の単語の辞書に切り換える(S503)。

【0076】また、その位置が認識対象の単語辞書を切り換える必要がない事を示している場合(NO)には、辞書を変更せずに、そのまま音声認識処理S502を実行する。

【0077】次に、音声認識処理S502において、認識結果として該当するものがあるか否かを示す判定する(S504)。入力した音声に対して、辞書の中に該当する単語がない場合(YES)には、図1に示した辞書切り換え部103は、音声認識部109から該当なしの認識結果110を受けて、次の候補の認識対象の単語辞書に切り換える(S505)。また、入力した音声に対して、辞書の中に該当する単語がある場合(NO)には、音声認識処理を終了し、認識結果に対してなされるシステムにおける次の処理へ移行する。

【0078】(3) カーナビゲーションシステムの応用例

次に、図5および図6を用いて本実施形態の音声認識システムのカーナビゲーションシステムのさらなる応用例について説明しよう。図5は、カーナビゲーションのエリアと対応する辞書の関係を、その地図上の構成物の位置関係を含めて説明するための模式図である。図6は、音声認識の結果により、カーナビゲーションシステムの地図上に結果を反映する処理を示す模式図である。図7

は、本実施形態のカーナビゲーションシステムのディスプレイの表示を示す模式図である。

【0079】ここで説明する応用例は、本発明の音声認識システムをカーナビゲーションシステムに適用し、さらに、使用者に対するユーザインタフェースを向上させるものである。

【0080】図5(a)に示されるように、カーナビゲーションシステムのディスプレイには、丸の記号で示した車が道路901を走行している様子が表示される。

【0081】記号902は、カーナビゲーションシステムを搭載した車が現在走行しているポイント(A地点)と走行方向を表示している。

【0082】A地点において、音声認識可能な単語は、904が示すエリア1の中に存在する地名、通り名、交差点名、建造物名、個人住宅(個人名)等であることは、既に説明した通りである。

【0083】また、この応用例での辞書のフォーマット908は、図5(b)に示す如くである。本実施形態では、エリア毎に辞書が対応しているので、エリアごとに単語をブロック化する必要がある。フィールド909は、そのための対応するエリアを示す番号である。フィールド910は、各エリアに登録されている単語群である。フィールド911は、各単語の示す場所の絶対位置座標であり、例えば、経度 x_i 、緯度 y_i に相当するものが、各単語に付加情報として登録されている。

【0084】フィールド912は、拡張用の付加情報である。例えば、ファミリーレストランであれば、その店の電話番号、FAX番号、営業日、営業時間、利用のためのメモ、店案内などカーナビゲーションシステムに利用するさまざまな情報を記憶しておけばよい。

【0085】矩形904が示すエリア1の辞書では、ガソリンスタンド(短縮形として、「ガソリン」)コンビニエンスストア(短縮形として、「コンビニ」)、ファミリーレストラン(短縮形として、「レストラン」)等が登録されており、単語「ガソリンスタンド」には、このエリア1に存在するガソリンスタンドの絶対位置座標値(x_i , y_i)が、位置情報として付加されている。同様に、単語「コンビニエンスストア」には、このエリア1に存在するコンビニエンスストアの絶対位置座標値(x_{i+1} , y_{i+1})が、位置情報として付加されている。同様に、単語「ファミリーレストラン」には、このエリア1に存在するファミリーレストランの絶対位置座標値(x_{i+2} , y_{i+2})が、位置情報として付加されている。また、905が示すエリア2の辞書では、郵便局、個人の住宅として鈴木宅、佐藤宅等が登録されており、単語「郵便局」には、このエリア2に存在する郵便局の絶対位置座標値(x_k , y_k)が、位置情報として付加されている。同様に、単語「鈴木宅」には、このエリア2に存在する個人住宅である鈴木宅の絶対位置座標値(x_{k+1} , y_{k+1})が、位置情報として付加されている。

単語「佐藤宅」には、このエリア2に存在する個人住宅である佐藤宅の絶対位置座標値 (x_{k+2} , y_{k+2}) が、位置情報として付加されている。

【0086】いま例えば、カーナビゲーションシステムのディスプレイに表示されている矩形904のエリア1において、走行中の運転者は、コンビニエンスストアに入りたいと思ったとする。そこで、カーナビゲーションシステムに対して、例えば、「コンビニ」と発声すると、エリア1内にコンビニエンスストアが906のC地点に存在すれば、まず、図7に示すように目的地C地点を点滅表示し、目的地C地点までのルートを実線や色を変えて表示し、走行距離や走行所要時間を表示する。また、音声誘導で「3 km先の次の交差点を左折して、次の交差点を右折したところにあります。」と音声合成で答えてくれる。

【0087】次に、図6を用いてこのような音声によるインタフェースを実現するための処理を、上の例により説明しよう。

【0088】先ず、音声認識システムは、各エリアごとの単語辞書に対して音声認識をおこなう (S1001)。認識された単語、例えば、「コンビニ」には、コンビニが存在する場所の絶対位置座標値である (x_{111} , y_{111}) が、認識結果 (目的地) として、「コンビニ」を示すテキストコードと共に出力される。

【0089】次に、S1001で出力された認識結果 (目的地)、「コンビニ」の絶対位置座標値 (x_{111} , y_{111}) から、カーナビゲーションシステムのディスプレイ等の表示デバイス座標系の座標値で、かつ表示されている方角による座標値に座標変換計算する (S1002)。

【0090】また、S1001で出力された認識結果 (目的地)、「コンビニ」の絶対位置座標値 (x_{111} , y_{111}) と、現在走行している現在地を示すGPSからの絶対位置座標値 (X_c , Y_c) とから、道路事情を含めた、現在地から目的地までの走行距離を計算する (S1003)。

【0091】最後に、上記の計算結果に基づいて、カーナビゲーションシステムのディスプレイ上に、目的地、目的地までのルート、走行距離、走行所要時間を表示する (S1004)。

【0092】次に、今一つのユーザインタフェースを提供する例について説明しよう。

【0093】さて、A地点を走行していた車は、現在、エリア2にあるB地点を走行しているものとする。

【0094】記号903は、カーナビゲーションシステムを搭載した車が現在走行しているポイント (B地点) と走行方向を表示している。B地点において、音声認識可能な単語は、905が示すエリア2の中に存在する地名、通り名、交差点名、建造物名、個人住宅 (個人名) 等である。

【0095】いま例えば、カーナビゲーションシステムのディスプレイに表示されている矩形905のエリア2において、走行中の運転者は、最終目的地である「佐藤宅」に向かっているとすると、そこで、カーナビゲーションシステムに対して、例えば、「佐藤宅」と発声すると、エリア2内に佐藤宅が907のD地点に存在すれば、まず、最終目的地D地点を点滅表示し、目的地D地点までのルートを実線や色を変えて表示し、走行距離や走行所要時間を表示する。また、音声誘導で「5 km先の次の信号機交差点を右折して、次の交差点を左折したところにあります。」と音声合成で答えてくれる。

【0096】上記の処理は、図4を用いて説明したようにエリア1の場合と同様に計算して、おこなうことができる。

【0097】〔実施形態2〕以下、本発明に係る第二の実施形態を、図8ないし図10を用いて説明する。図8は、携帯型情報機器の概観図である。図9は、本発明に係る第二の実施形態の音声認識システムの辞書の切り換え処理と音声認識の処理を説明するフローチャートである。図10は、音声認識のための辞書の階層を図示した模式図である。

【0098】本実施形態は、本発明の音声認識システムを、PDA (Personal Digital Assistants) に代表されるような携帯型情報機器、携帯型翻訳機等のシステムに、応用した場合である。

【0099】図8に示されるような携帯型情報機器は、半導体技術の進歩により、年々小型で便利なものが開発されており、昨今では、爆発的な普及を見ている。本発明の音声認識のために辞書を切り換えるというアイデアは、このような携帯型情報機器にも利用することができる。

【0100】すなわち、図10に示されるように、音声認識のためのインデックス辞書を用意しておく。このインデックス辞書には、例えば、「開け」、「保存」、「印刷」などのコマンド、「住所録」、「スケジュール」、「メモ」などのこの機器で利用できる機能が登録されている。

【0101】そして、このインデックス辞書に該当する語が認識されたときには、それに関連する文類別辞書に切り換えるようにする。例えば、「住所録」の語が、認識されたときには、その住所録機能のための人名が登録された辞書である。また、「開け」コマンドに関する辞書は、このコマンドに対するオプションのための語を登録した辞書である。

【0102】この処理を図9の順を追って説明すると以下の通りである。

【0103】先ず、音声認識のための辞書を音声認識の辞書として使える状態にして、この辞書に対して音声認識させる (S601)。

【0104】次に、インデックス辞書の認識結果に対し

て、認識結果が示す辞書に切り換える（S602）。

【0105】最後に、新たに音声認識の対象になった分類別辞書により、音声認識がおこなわれる（S603）。例えば、上の例で言うと、「日立太郎」と音声入力すると、「日立太郎」を音声認識処理して、その機器のディスプレイに日立太郎の住所が出力される。

【0106】

【発明の効果】本発明によれば、小型情報システムに用いられる音声認識システムにおいて、システム全体として音声認識できる語彙数の増加しても、認識率や認識応答時間の性能を低下させないで音声認識ができる音声認識システムを提供することができる。

【0107】また、本発明によれば、音声認識を用いたカーナビゲーションシステムにおいて、特に、音声による目的地の地図上の位置の探索や、音声による誘導システムにおいて、良好な音声認識インターフェースを実現することができる。

【図面の簡単な説明】

【図1】本発明に係る音声認識システムの各機能とその処理の流れを示すブロック図である。

【図2】本発明のハードウェア構成を示すブロック図である。

【図3】カーナビゲーションのエリアと対応する辞書の関係を説明するための模式図である。

【図4】本発明の第一の実施形態に係る音声認識システムにおいて、音声認識の辞書切り換え処理を示すフローチャートである。

【図5】カーナビゲーションのエリアと対応する辞書の＊

＊関係を、その地図上の構成物の位置関係を含めて説明するための模式図である。

【図6】音声認識の結果により、カーナビゲーションシステムの地図上に結果を反映する処理を示す模式図である。

【図7】本実施形態のカーナビゲーションシステムのディスプレイの表示を示す模式図である。

【図8】携帯型情報機器の概観図である。

【図9】本発明に係る第二の実施形態の音声認識システムの辞書の切り換え処理と音声認識の処理を説明するフローチャートである。

【図10】音声認識のための辞書の階層を図示した模式図である。

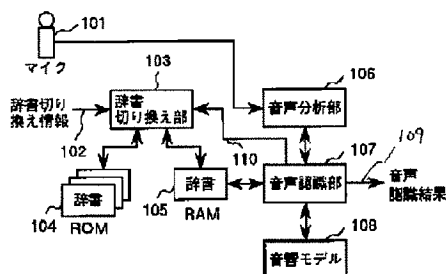
【図11】従来技術に係る音声翻訳装置の構成を示すブロック図である。

【符号の説明】

901…カーナビゲーションシステム道路地図表示
902…カーナビゲーションシステムA地点走行車表示
903…カーナビゲーションシステムB地点走行車表示
904…カーナビゲーションシステムエリア1表示
905…カーナビゲーションシステムエリア2表示
906…カーナビゲーションシステム目的地C地点表示
907…カーナビゲーションシステム目的地D地点表示
908…カーナビ用音声認識単語辞書フォーマット
909…カーナビ用音声認識単語辞書エリア
910…カーナビ用音声認識単語辞書単語
911…カーナビ用音声認識単語辞書位置座標値
912…カーナビ用音声認識単語辞書拡張子

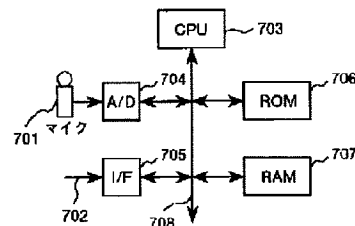
【図1】

図 1



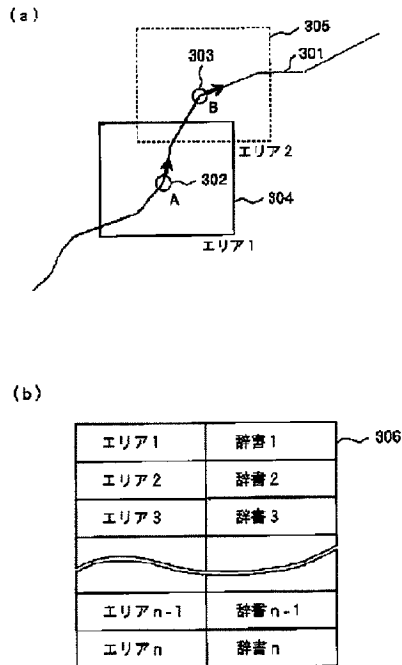
【図2】

図 2



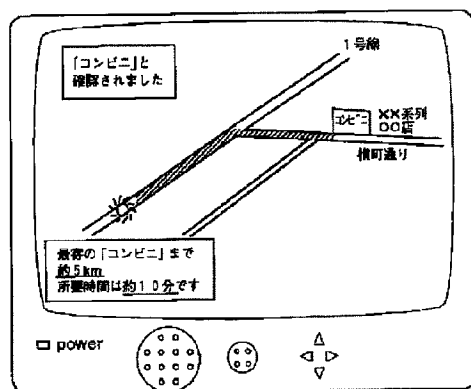
【図3】

図 3



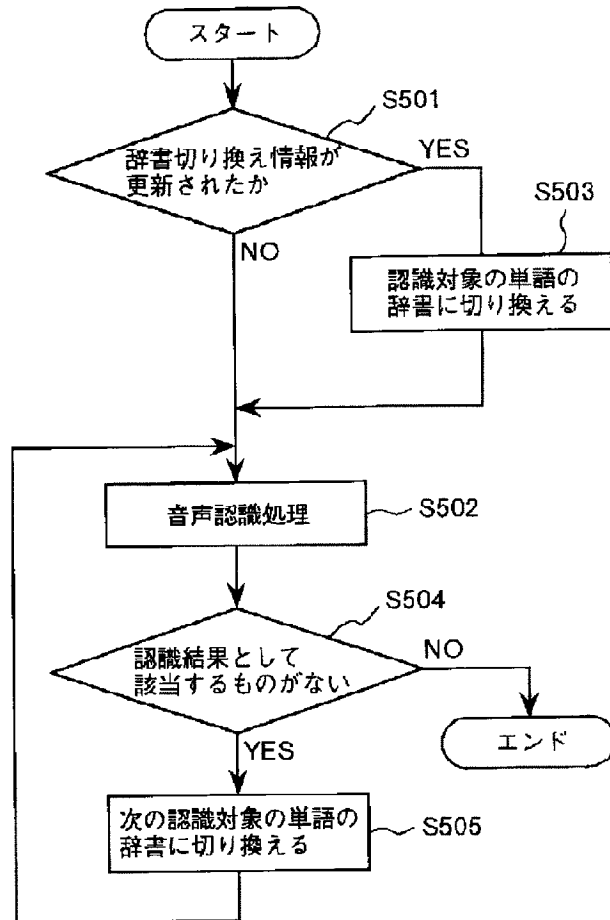
【図7】

図 7



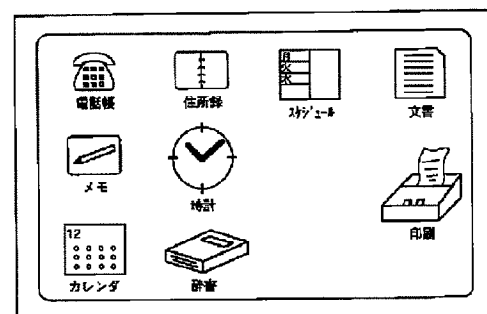
【図4】

図 4



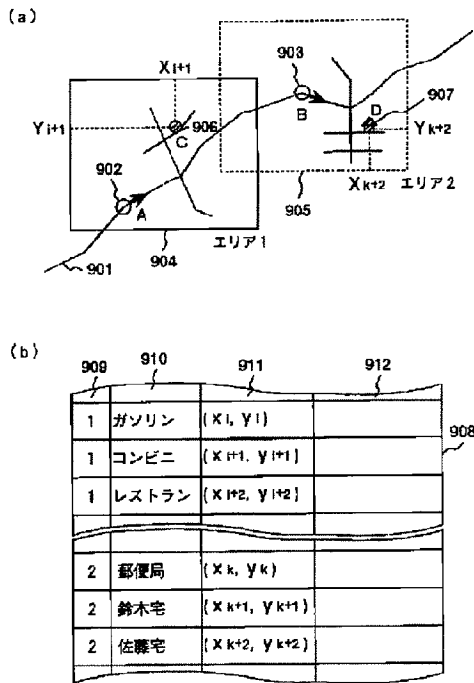
【図8】

図 8



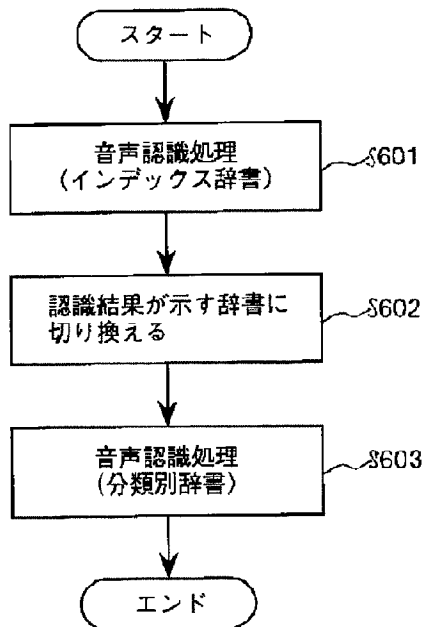
【図5】

図 5



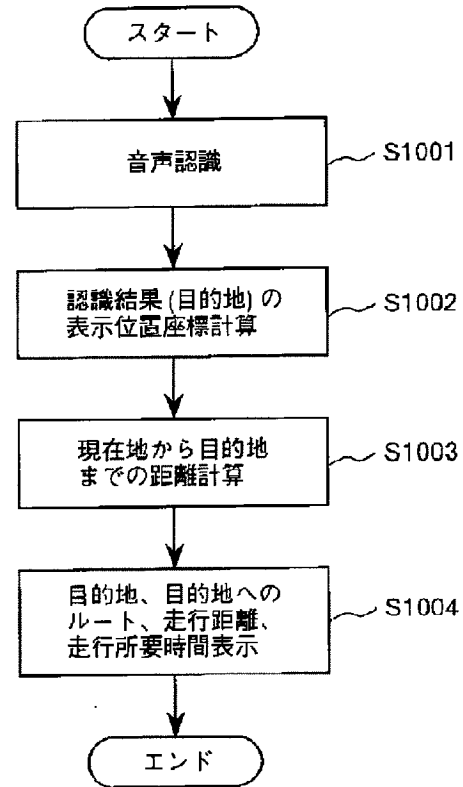
【図9】

図 9



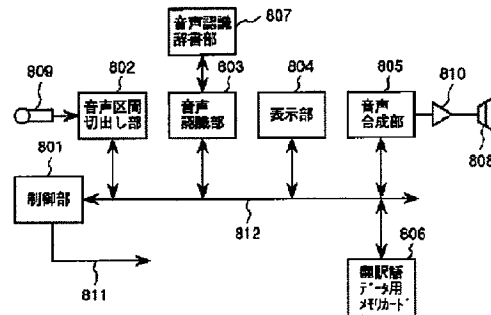
【図6】

図 6



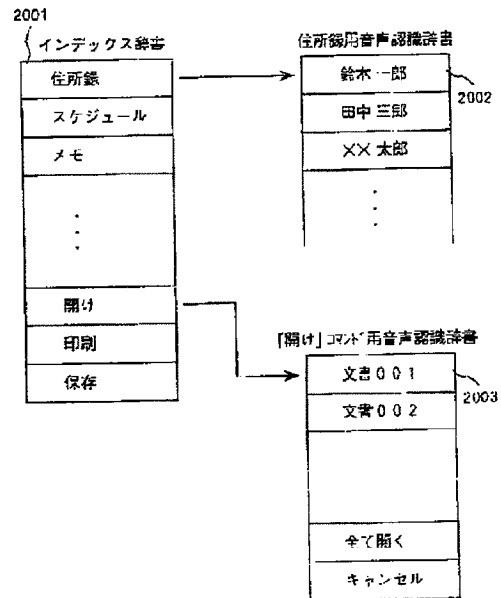
【図11】

図 11



【図10】

図 10



フロントページの続き

(51)Int.Cl.⁶
G 0 9 B 29/10

識別記号

F I
G 0 9 B 29/10

A